



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 02:58 PM GMT

PDB ID : 4HD7  
Title : Crystal Structure of Tyrosinase from Bacillus megaterium V218G mutant soaked in CuSO<sub>4</sub>  
Authors : Goldfeder, M.; Kanteev, M.; Adir, N.; Fishman, A.  
Deposited on : 2012-10-02  
Resolution : 2.10 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

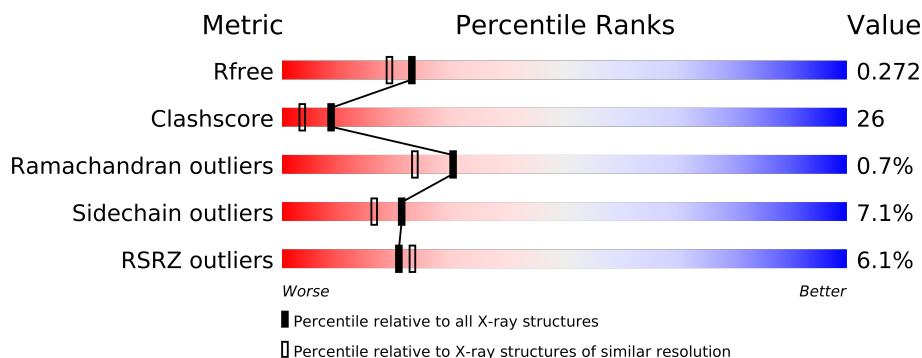
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	303	
1	B	303	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4914 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Tyrosinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	0	1	0
			2353	1495	426	424	8			
1	B	286	Total	C	N	O	S	0	0	0
			2343	1489	423	423	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	SER	ENGINEERED MUTATION	UNP B2ZB02
A	218	GLY	VAL	ENGINEERED MUTATION	UNP B2ZB02
A	298	HIS	-	EXPRESSION TAG	UNP B2ZB02
A	299	HIS	-	EXPRESSION TAG	UNP B2ZB02
A	300	HIS	-	EXPRESSION TAG	UNP B2ZB02
A	301	HIS	-	EXPRESSION TAG	UNP B2ZB02
A	302	HIS	-	EXPRESSION TAG	UNP B2ZB02
A	303	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	2	GLY	SER	ENGINEERED MUTATION	UNP B2ZB02
B	218	GLY	VAL	ENGINEERED MUTATION	UNP B2ZB02
B	298	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	299	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	300	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	301	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	302	HIS	-	EXPRESSION TAG	UNP B2ZB02
B	303	HIS	-	EXPRESSION TAG	UNP B2ZB02

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Cu	0	0
			2	2		
2	A	2	Total	Cu	0	0
			2	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	109	Total 109	O 109	0	0
3	B	105	Total 105	O 105	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.90Å 78.67Å 85.77Å 90.00° 105.91° 90.00°	Depositor
Resolution (Å)	36.21 – 2.10 41.24 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.4 (36.21-2.10) 95.4 (41.24-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.53 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.254 , 0.282 0.246 , 0.272	Depositor DCC
$R_{free}$ test set	3424 reflections (10.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	19.2	Xtriage
Anisotropy	0.451	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 28.5	EDS
Estimated twinning fraction	0.039 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 34243 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	4914	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.21% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CU

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.48	0/2432	0.59	0/3313
1	B	0.43	0/2421	0.61	0/3298
All	All	0.46	0/4853	0.60	0/6611

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2353	0	2227	106	1
1	B	2343	0	2221	131	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	109	0	0	28	0
3	B	105	0	0	34	0
All	All	4914	0	4448	235	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 26.

The worst 5 of 235 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:30:LYS:HE2	3:A:589:HOH:O	1.22	1.35
1:B:288:ILE:H	1:B:289:GLU:CA	1.54	1.20
1:B:288:ILE:N	1:B:289:GLU:HA	1.59	1.15
1:A:81:GLN:HB2	3:A:566:HOH:O	1.44	1.15
1:A:23:VAL:N	3:A:569:HOH:O	1.78	1.14

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:101:GLN:CG	1:B:101:GLN:NE2[1_454]	1.95	0.25

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	285/303 (94%)	275 (96%)	10 (4%)	0	100	100
1	B	284/303 (94%)	266 (94%)	14 (5%)	4 (1%)	16	9
All	All	569/606 (94%)	541 (95%)	24 (4%)	4 (1%)	30	23

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	288	ILE
1	B	203	LEU
1	B	245	HIS
1	B	244	ILE

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain



conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	249/264 (94%)	232 (93%)	17 (7%)	22	18
1	B	248/264 (94%)	230 (93%)	18 (7%)	20	15
All	All	497/528 (94%)	462 (93%)	35 (7%)	21	17

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	281	LYS
1	B	54	SER
1	B	247	ASN
1	A	288	ILE
1	B	4	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	270	ASN
1	B	117	ASN
1	B	261	ASN
1	B	57	ASN
1	B	84	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	286/303 (94%)	0.53	11 (3%)	38 43	11, 21, 45, 82	1 (0%)
1	B	286/303 (94%)	0.89	24 (8%)	11 12	13, 26, 50, 86	1 (0%)
All	All	572/606 (94%)	0.71	35 (6%)	21 23	11, 23, 49, 86	2 (0%)

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	247	ASN	9.3
1	B	246	ARG	7.1
1	B	5	TYR	6.9
1	B	248	GLN	6.8
1	A	5	TYR	6.2

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CU	A	402	1/1	0.10	-2.76	48,48,48,48	1
2	CU	B	402	1/1	0.10	-2.82	48,48,48,48	1
2	CU	B	401	1/1	0.05	-4.94	34,34,34,34	1
2	CU	A	401	1/1	0.04	-7.62	26,26,26,26	0

## 6.5 Other polymers

There are no such residues in this entry.